geoconvention
Calgary • Canada • May 12-14 2025

# From Imagery to Graphs with Scalable AI Driven Pedestrian Pathway Mapping

*Yuxiang Zhang[1], Suresh Devalapalli[1], Ravi Gunti[1], Anat Caspi[2]*
*Gaussian Solutions LLC[1], University of Washington[2]*

## Summary

Advances in geoscience and AI systems have made it possible to automatically generate large-scale pedestrian pathway datasets, overcoming challenges that previously hindered their creation. Pedestrian networks, unlike road networks, require detailed representations of connectivity, transitions, and attributes necessary for urban planning and geospatial analysis. However, existing datasets are often sparse, unreliable, and non-interoperable due to fragmented collection methods and jurisdictional variability [1, 2].

Proposed solutions to create pathway datasets involve some combination of direct inspection of the built environment, computational inference of remotely sensed data, or interactive review of existing sources. Direct inspection can provide highly detailed and accurate data with sufficient training but does not scale due to prohibitive costs. Computational inference scales broadly but struggles with unusual situations and occlusions, leading to biased outputs [3]. Interactive repair combines automated and manual approaches, allowing for modest-scale data collection and quality control, but remains sensitive to the training and experience of the operators. We suggest that all three approaches, organized interdependently, offer the only viable solution.

Prior work has recognized the need for combining multiple sources of information. For example, Wu et al. [4] and Sun et al. [5] demonstrated the utility of integrating OpenStreetMap (OSM) data, GPS, and satellite imagery for road extraction. Recent advancements in remote sensing have led to methods to automate pedestrian pathway data collection. For example, Ning et al. [6] applied neural networks to extract sidewalks from aerial images, and Tile2Net [7] inferred pedestrian pathway features from satellite imagery. While these studies have improved pedestrian environment mapping, they do not generate a comprehensive, connected, and routable pedestrian pathway graph necessary for city planning and navigation. Our methodology builds on these foundations, emphasizing the importance of connected, routable pedestrian pathways.

In this work, we introduce an innovative AI-driven end-to-end pipeline for the scalable collection, processing, and maintenance of pedestrian pathway data (Figure 1). By integrating multi-input segmentation models with aerial imagery and existing road network datasets, we extract features such as sidewalks, crossings, and curb ramps and create connected, routable pedestrian pathway graphs (Prophet, Figure 1 left), ensuring compatibility with broader geospatial datasets and enabling their use in urban planning and mobility solutions. By organizing predictions into project regions and subdividing them into manageable tasks, our hybrid workflow ensures efficiency and quality. Our interactive toolsets (Skeptic, Figure 1 right) tracks progress, manage data consistency, and incorporate user feedback to iteratively improve the network quality.

Our proposed system significantly outperforms state-of-the-art methods, reducing the need for extensive human intervention and enabling scalable implementation across large areas. This

work demonstrates the transformative potential of AI in geospatial sciences, offering robust solutions for high-quality pedestrian network data generation. By addressing technical and organizational challenges, our approach represents a significant advancement in applying geoscience and AI to support urban planning and mobility technologies.

## Method

*Prophet: Network-Centered Inference of Pedestrian Pathway Networks.* We developed a fully automated method for pedestrian pathway graph inference called Prophet. As shown in Figure 1 (left), the process consists of three main steps. First, we developed a module to create a hypothesized graph, by inferring an optimistic sidewalk network from existing street networks, inferring street crossing locations, and creating a preliminary conjecture of curb interface locations. Second, we use a multi-input semantic segmentation network to generate pixel-wise prediction masks for the important classes in the pedestrian environment, utilizing both satellite images and rasterized existing street information. Lastly, we use information from the semantic segmentation network to optimize the hypothesized graph into an accurate connected pedestrian pathway graph.

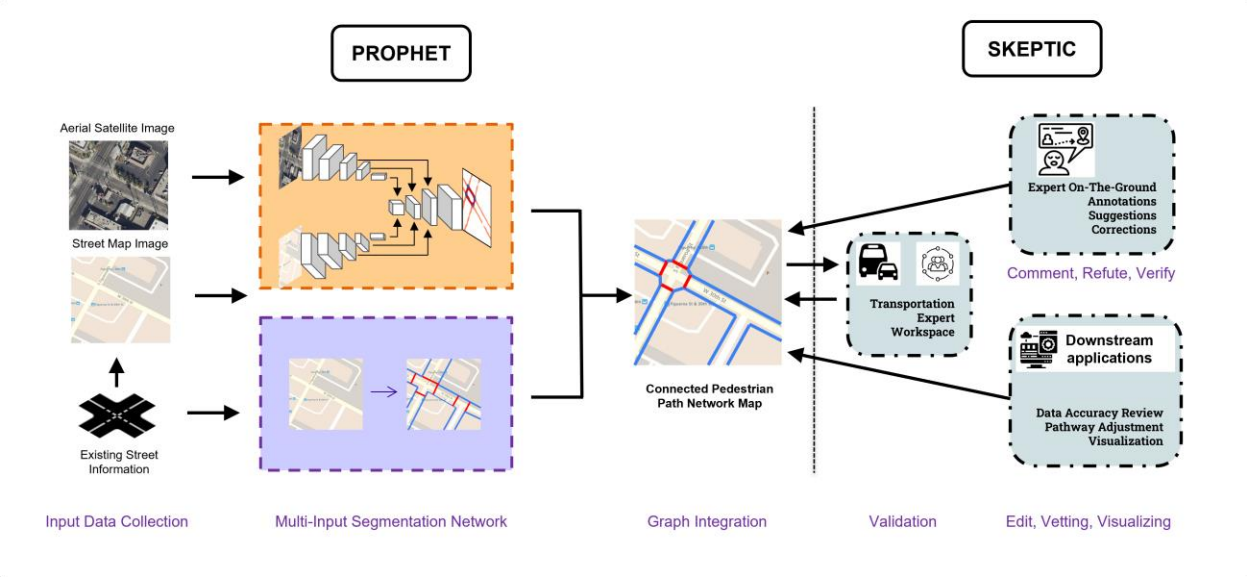



Figure 1: Overview of our end-to-end process for inferring pedestrian pathway graph. Left: Prophet, the multi-input framework that automates the generation of pedestrian path networks. Right: The Skeptic protocol and toolsets for manual vetting and community engagement.

*Skeptic: Collaborative Edits, Validation, and Verification.* To provide a scalable and standardized tool to review and validate the AI generated pathway graph, we developed Skeptic (Figure 1 right), which provides a set of instructions and collaborative tooling to engage stakeholders in expert review and validation of the predicted pedestrian pathway graph. The

Skeptic protocol was designed in response to the findings from our interviews with professionals working at thirteen municipal-, county- and state-level governments in Maryland, Oregon, and Washington States. Skeptic prioritizes diverse regions and uses intersection-scale tasks to streamline validation and enhance collaboration. A modified version of the Humanitarian OpenStreetMap Team (HOT) [8] tasking tool is employed to manage edits, track progress, and ensure interoperability with OSM tools. Volunteers validate and enrich the data with local details, supporting accessible routing and future model training.

## Results

Figure 2 visualizes the intermediate segmentation results and the final connected pathway graphs produced by Prophet. These examples highlight the difficulty in predicting pedestrian path network with single-source input and the improvement gained from adding other input sources.

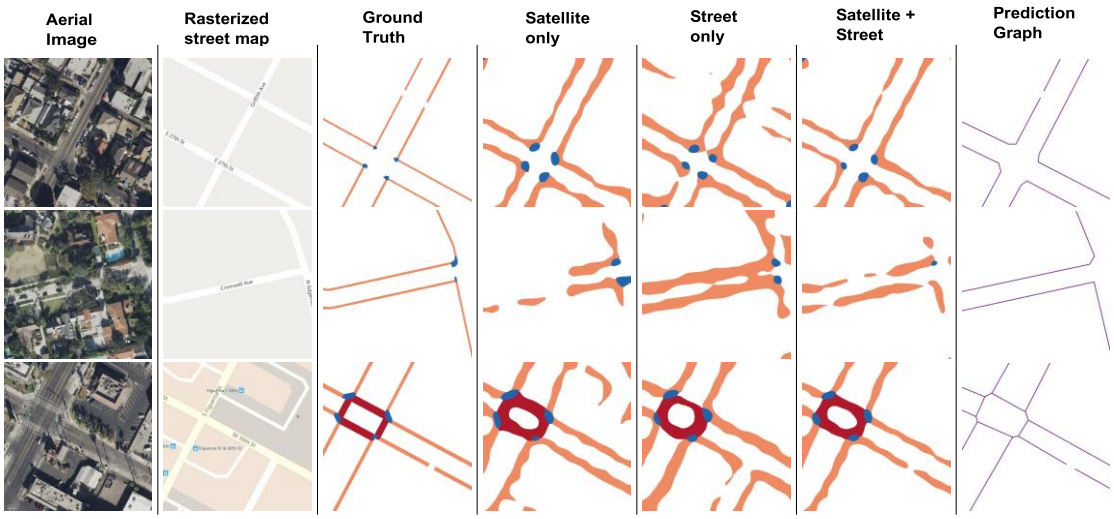



Figure 2: Visualization of the results. Columns 1-2 show the inputs to our method. The segmentation results of 3 different models are shown in columns 4-6. (1) Trained with the aerial satellite image only (2) Trained with the street map tile only (3) Trained with both the aerial satellite image and the street map image tile. The segmented classes are *Corner bulb*, *Sidewalk*, and *Crossing*. The model using both aerial satellite images and street map images outperforms models that use either input alone. Column 7 illustrates the connected pedestrian pathway graph inferred by Prophet.

To measure routability, we use graph-level metrics proposed in previous work [9]. Three categories of metrics are considered (1) local, polygon-level routability metrics including average number of connected components (avg CC), the average betweenness centrality (avg

BC), and *TraversabilitySimilarity*, (2) count-based metrics (node count, edge count, and avg degree for the entire test area), and (3) the F1 scored based on the network segment edge-retrieval method described in previous work [3]. *TraversabilitySimilarity* is a specialized metric designed to provide a local estimate of global routability and enable comparisons across predicted networks from different sources [9], it rewards predictions that reproduce connected paths for traversing an intersection.

Evaluation results are summarized in Table 1. The metrics in the 'Global' and 'Local' columns indicate whether the property applies to the overall graph (global) or the average of a polygon-local property (local). A predicted graph with values closer to the Ground Truth indicates better model performance. The 'Edge-retrieval F1' and *'TraversabilitySimilarity'* metrics provide direct comparisons to the Ground Truth, where higher scores indicate better model performance. In particular, *TraversabilitySimilarity* compares whether local intersections afford the same ingress and egress points in both predicted and ground truth graphs. Avg CC and avg BC provide additional measures of local connectivity, where the distance from ground truth values exposes deviations from the ground truth graph structure. F1 score captures similarity between predicted and ground truth graphs in terms of single edge's accuracy, but it does not reflect how travelers navigate the environment. The F1 score is therefore less useful for model evaluation and model selection in downstream applications that rely on routing. On the other hand, *TraversabilitySimilarity* better aligns with the practical requirements of real-world routing and navigation tasks. A higher *TraversabilitySimilarity* score of a model indicates a more accurate representation of the true connectivity and routability in its generated graph, as compared to the Ground Truth graph.

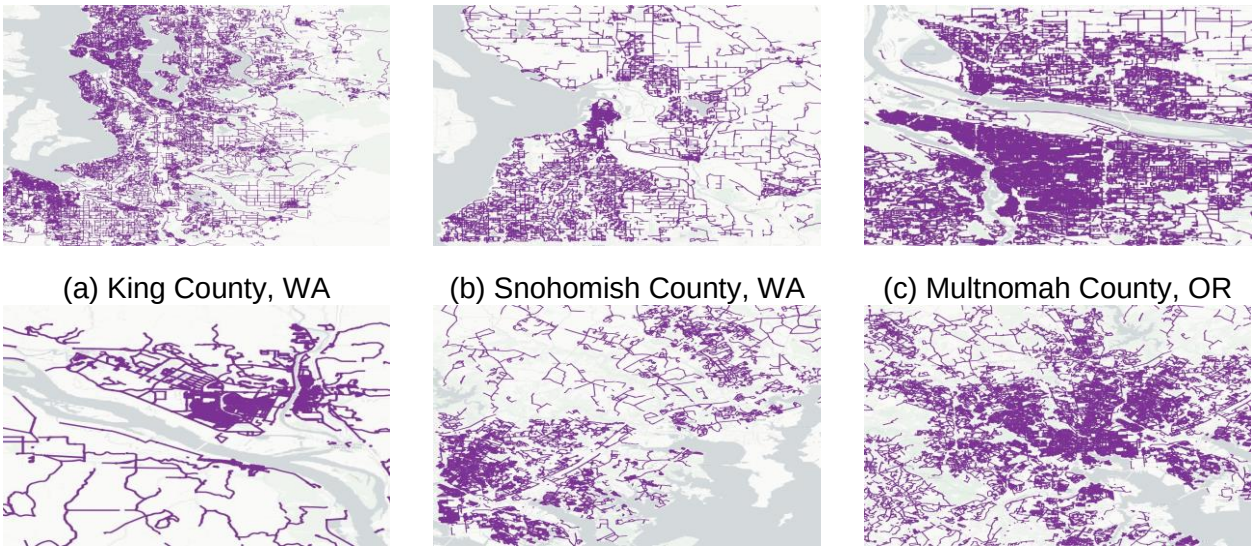
TABLE 1: Pathway Network Graph Routability Evaluation

| Method | Area | Global | | | Local | | Local (relative to Ground Truth) | |
|---|---|---|---|---|---|---|---|---|
| | | # nodes | # edges | avg degree | avg CC | avg BC | edge-retrieval F1 | *TraversabilitySimilarity* |
| Ground Truth | Argyle | 7576 | 8331 | 0.17 | 1.92 | 0.12 | 1.0 | 1.0 |
| | Bertha Pitts | 6994 | 7693 | 0.17 | 2.07 | 0.11 | 1.0 | 1.0 |
| | Burbridge | 3948 | 4408 | 0.20 | 1.82 | 0.12 | 1.0 | 1.0 |
| Tile2net | Argyle | 8386 | 7872 | 0.18 | 3.44 | 0.05 | 0.98 | 0.32 |
| | Bertha Pitts | 7744 | 7274 | 0.16 | 3.60 | 0.04 | 0.98 | 0.39 |
| | Burbridge | 4349 | 3954 | 0.24 | 3.13 | 0.04 | 0.98 | 0.31 |
| Prophet | Argyle | 3623 | 3932 | 0.32 | 1.40 | 0.11 | 0.97 | 0.47 |
| | Bertha Pitts | 2964 | 3169 | 0.34 | 1.25 | 0.10 | 0.97 | 0.44 |
| | Burbridge | 2383 | 2630 | 0.26 | 0.88 | 0.13 | 0.96 | 0.31 |
| Prophet + Skeptic | Argyle | 4399 | 5019 | 0.12 | 1.76 | 0.11 | 0.99 | **0.84** |
| | Bertha Pitts | 3424 | 3885 | 0.19 | 1.87 | 0.10 | 0.99 | **0.88** |
| | Burbridge | 3137 | 3489 | 0.19 | 1.85 | 0.11 | 0.99 | **0.84** |

Three methods are included in the evaluation: (1) **Tile2Net** [7]: This state-of-the-art method uses aerial satellite images to segment sidewalks, crosswalks, and footpaths in cities. It then

simplifies the segmented polygons and extracts the centerlines of the polygons to represent pathway graphs. (2) **Prophet**: Our method for generating pathway graphs, using both the aerial satellite images and street map image as input and the ViT-Base [10] as backbone in the segmentation model. (3**) Prophet + Skeptic**: Facilitated human editing applied to Prophet outputs.

As shown in Table 1, our methods (Prophet, and Prophet+Skeptic) outperform the state-of-the-art method (Tile2Net) in terms of capturing graph traversability and connectivity as compared to the ground truth. With the addition of Skeptic, mappers can start from the pathway graphs generated by Prophet, and make edits as needed efficiently (because most of the graph components are already populated by Prophet). Using Prophet and Skeptic together generates graphs that accurately capture the true connectivity and routability in the built environment (above 0.8 *TraversabilitySimilarity* score in all test areas.



(a) King County, WA (b) Snohomish County, WA (c) Multnomah County, OR

(d) Columbia County, OR (e) Harford County, MD (f) Baltimore County, MD

FIGURE 3: Pedestrian pathway graphs data generated with Prophet

TABLE 2: Statistics on data generated with Prophet for U.S. counties

| Area | Number of sidewalks | Length of sidewalks (m) | Number of crossings | length of crossings (m) | Number of intersections |
|---|---|---|---|---|---|
| King | 249683 | 15850008 | 133603 | 1968246 | 45971 |
| Snohomish | 62991 | 5033712 | 33483 | 480327 | 12388 |
| Multnomah | 178631 | 10647695 | 92788 | 1385683 | 31697 |
| Columbia | 17762 | 2253858 | 9291 | 139957 | 3231 |
| Harford | 60979 | 5141570 | 32224 | 460859 | 11110 |
| Baltimore | 159842 | 11379123 | 84479 | 1215692 | 28432 |

TABLE 3: Statistics on data generated with Prophet+Skeptic

| State | Area (km2) | Sidewalk Count | Curb Count | Crossing Count | Sidewalk Length (km) | Crossings Length (km) |
|---|---|---|---|---|---|---|
| Washington | 6332.35 | 467401 | 633963 | 452083 | 33635.32 | 5720.4 |

After demonstrating the robust performance of Prophet in multiple test areas, we used Prophet to generate pedestrian pathway graphs at scale. Specifically, we generated connected pedestrian pathway graphs for 6 counties in the U.S., including (1) King County, WA (2) Snohomish County, WA (3) Multnomah County, OR (4) Columbia County, OR (5) Harford County, MD (6) Baltimore County, MD. Figure 3 illustrates the prediction graphs, and Table 2 shows their aggregate statistics. These graphs generated by Prophet can fill the information gap in places where no other routable pathway graphs are available and can be refined with Skeptic for more accurate representation for the actual pedestrian pathway network. Table 3 shows the aggregate statistics of the data produced with (Prophet + Skeptic) for Washington State. The data can be viewed at viewer.tdei.us.

## Conclusions and Future Work

In this work, we introduce a sociotechnical protocol combining automated inference with manual review and community engagement to achieve routable, reliable, and reproducible pedestrian pathway graphs. We piloted this protocol in WA State, demonstrating that Prophet outperforms imagery-only state-of-the-art methods, but that manual review with Skeptic offers further improvements. We described a community engagement protocol to further enrich and validate the data for the needs of those most affected by quality issues. In ongoing work, we are implementing feedback loops into the process to allow vetted data to be used to continuously train and improve the model and design redundant review to verify inter-mapper consistency.

## Acknowledgements

This work was funded in part by the Taskar Center for Accessible Technology, USDOT ITS4US NOFO No: 693JJ322NF00001, and Microsoft's AI4Accessibility award.

### References

1. Bolten, N. and A. Caspi, Towards routine, city-scale accessibility metrics: Graph theoretic interpretations of pedestrian access using personalized pedestrian network analysis. PLoS one, Vol. 16, No. 3, 2021, p. e0248399.

2. Bolten, N. and A. Caspi, Towards operationalizing the communal production and management of public (open) data: a pedestrian network case study: A pedestrian network case study in operationalizing communal open data. In ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies (COMPASS), 2022, pp.

232–247.

3. Rhoads, D., C. Rames, A. Solé-Ribalta, M. C. González, M. Szell, and J. BorgeHolthoefer, Sidewalk networks: Review and outlook. Computers, Environment and Urban Systems, Vol. 106, 2023, p. 102031.

4. Wu, S., C. Du, H. Chen, Y. Xu, N. Guo, and N. Jing, Road extraction from very high resolution images using weakly labeled OpenStreetMap centerline. ISPRS International Journal of Geo-Information, Vol. 8, No. 11, 2019, p. 478.

5. Sun, T., Z. Di, P. Che, C. Liu, and Y. Wang, Leveraging crowdsourced GPS data for road extraction from aerial imagery. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 7509–7518.

6. Ning, H., X. Ye, Z. Chen, T. Liu, and T. Cao, Sidewalk extraction using aerial and street view images. Environment and Planning B: Urban Analytics and City Science, Vol. 49, No. 1, 2022, pp. 7–22.

7. Hosseini, M., A. Sevtsuk, F. Miranda, R. M. Cesar Jr, and C. T. Silva, Mapping the walk: A scalable computer vision approach for generating sidewalk network datasets from aerial imagery. Computers, Environment and Urban Systems, Vol. 101, 2023, p. 101950.

8. HOT Tasking Manager, HOT Tasking Manager. https://tasks.hotosm.org/, retrieved July 26, 2024.

9. Zhang, Y., B. Howe, S. Mehta, N.-J. Bolten, and A. Caspi, PathwayBench: Assessing Routability of Pedestrian Pathway Networks Inferred from Multi-City Imagery. arXiv preprint arXiv:2407.16875, 2024.

10. DosoViTskiy, A., L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. arXiv preprint arXiv:2010.11929, 2020.